

chapter 3

THE CHEMICAL NATURE OF GENETIC MATERIAL

There can be little doubt that the major share of the heredity of a strain is carried in the chromosomes of its germ cells. This conclusion is based firmly on the observations of the cytologist and the geneticist that specific mutations may be directly related to localized morphological changes in chromosomes. It serves as the starting point for one of the most thriving enterprises in modern biological research, namely, the identification and chemical description of genetic material. Progress has been rapid and we can now state with some assurance that the substance most directly associated with the storage and perpetuation of hereditary information is the deoxyribonucleic acid (DNA) of the chromosomal strands.

Aside from the circumstantial evidence furnished by the cytochemical localization of DNA in chromosomes, there are a number of other lines of evidence which give more explicit information bearing on this idea. It was demonstrated by Boivin, Vendrely, and Vendrely¹ in 1948, for example, that the DNA content of somatic cells (diploid) was constant from tissue to tissue in a single species but that sperm cells (haploid) contained exactly half as much.

These observations were extended to a number of other species by Mirsky² and his collaborators. The latter group also showed that the distribution and quantities of various other chemical components of nuclei generally *did not* correlate in a way to be expected for genetically critical material. We know that the nucleus contains, in addition to DNA, ribonucleoprotein, various arginine-rich protamines and histones, a tryptophan-rich protein fraction, and a small amount of lipid. None of these substances (with the exception of the protamine-histone fraction, which may be directly associated with DNA) appears to have the constancy of distribution from cell type to cell type within a species of the sort exhibited by the DNA component.

Another type of experimental observation which suggests a genetic role for DNA is concerned with the effects of mutagenic agents, such as ultraviolet radiation and certain chemicals. It has been shown that the efficiency spectrum of ultraviolet light in producing mutations is closely related to the absorption spectrum of nucleic acid. Such experiments are not completely convincing in themselves since the absorption of ultraviolet photons by the nucleic acid molecule might conceivably be only the first step in a chain of reactions in which the final target could reside in molecules of a rather different chemical nature. Experiments on the mutagenic effect of such agents as mustard gas are similarly inconclusive since, although the nucleic acids do appear to be much more chemically reactive to these substances *in vitro* than are the proteins, we cannot discount the special sensitivity of a particularly important member of the latter class of compounds. In spite of these possible objections, there is a certain amount of direct evidence which indicates that when we tamper with the chemical structure of DNA, be it by radiation or by chemical techniques, the rate of mutation is increased. Zamenhof³ and his colleagues have shown that when the thymine analogue, 5-bromouracil, is incorporated by a cell into the structure of its DNA, the frequency of mutation is greatly increased. The reasons for this stimulation of the mutation rate may be related to the presence of the abnormal pyrimidine base in the polynucleotide sequence of the DNA molecule, or it may equally well be the result of some aberration in the kinetics of the biosynthesis of DNA. Whatever the reason, we may at least conclude that the DNA molecule is implicated in this chemically induced mutagenesis.

The metabolic stability of DNA also supports the close association of DNA with genes. In spite of a number of early misleading reports, the consensus now clearly indicates that the DNA content of chromosomes does not change during the stages of division, nor do the

subcomponents of its structure undergo equilibration with extranuclear sources of DNA precursors. The most clear-cut support for this conclusion comes from cytochemical and radiochemical studies on the nuclei and chromosomes of growing tissues, particularly of plants. Howard and Pelc,⁴ for example, attacked the problem by growing roots of the English broad bean, *Vicia faba*, in the presence of radioactive orthophosphate. Autoradiographs of squash preparations of the root tissue were then prepared by the stripping film technique. An analysis of the way in which radioactivity was associated with the nuclei in various stages of division and in the resting stage indicated that incorporation of the isotope takes place only in the resting, interphase nucleus and that prophase and metaphase nuclei are not actively synthesizing nucleic acid. Autoradiographs of root tips which had been incubated with isotope for more extended periods permitted the further conclusion that the tagged nucleic acids of the chromosomes are passed on to daughter cells without intermediate degradation and resynthesis. These experiments, and the resulting conclusions, have been greatly refined, both as the result of improvements in the technique of chromosome autoradiography and by the study of purified isotopically labeled DNA from various sources. We shall return to these more recent studies after we have first considered the chemistry of DNA and the organization of the chromosome in more detail. However, the experiments of Howard and Pelc, even without embellishment, are quite satisfying from the genetic point of view since they suggest conservation of DNA and the physical continuity of the gene during cell replication.

Some of the best evidence for the central importance of DNA in genetics comes from studies on "transformation." This phenomenon, discovered in pneumococcus by F. Griffith in 1928, has since been extended to a number of other microorganisms. It involves the change in the genotype of one strain of cells that is produced by exposure to extracts of cells of a different strain.

Pneumococci generally exist in two forms. One forms "smooth" colonies on agar plates, possesses a type-specific polysaccharide capsular substance, and is virulent. The other forms "rough" colonies and lacks both the virulence and the polysaccharide of the smooth form. The smooth forms are genetically stable, and a strain characterized by a Type II polysaccharide, for example, does not spontaneously mutate to Type III. "Smooth" organisms do, however, mutate to rough forms, and this conversion appears to be irreversible. Griffith showed that when mice were subjected to mixed

injections of living non-encapsulated "rough" organisms and killed "smooth" organisms, living encapsulated bacteria could be isolated from the animal. The *progeny* of these transformed bacteria were also encapsulated, and the specific polysaccharide was shown to persist indefinitely through successive generations until spontaneous mutation to a rough form occurred.

Avery and his collaborators,⁵ and later Hotchkiss, Zamenhof, and others, have shown that the substance in Griffith's extracts which is responsible for the transformation has the chemical characteristics of DNA.⁶ The evidence for this is now very convincing, and the transformation of organisms for a host of genetic markers in addition to that controlling encapsulation can now be attributed, with relative certainty, to the DNA molecule. The transformation is definitely not caused by a protein contaminant.

Many of the DNA-borne characters were originally selected by exposure of bacteria to sink-or-swim situations (Figure 21). For example, when pneumococci are grown in the presence of streptomycin, essentially all the cells are killed by the drug. A very few cells, however, survive and continue to multiply, producing a culture which is resistant to the levels of streptomycin employed. These organisms have acquired, by a chance mutation, the physiological characteristics which permit them to occupy a new "ecological niche" in nature (although a rather unnatural one). Not only do the cells which survive streptomycin behave as a new and constant phenotype, but samples of DNA prepared from them possess the ability to transform other cells to a state of streptomycin resistance.

The genetic stability of transformable traits is highly suggestive of true gene transfer. It is difficult, however, to prove unequivocally that actual chromosomal material is being transferred by this process from cell to cell. The chemistry and cytology of bacterial nuclei is still quite obscure and, to make matters more difficult, genetic analysis of the sort that can be done with sweet peas or *Drosophila* is not easy with bacteria because multiplication takes place most commonly by mother-daughter division rather than by sexual mating. Sexual mating does take place fairly frequently in some microorganisms, and it seems likely that the transmission of transformed characters as unit particles of heredity could be studied by some direct hybridization technique if the proper choice of bacterium and other transformable traits were combined in one experimental system. To my knowledge, however, such a study has not been carried out for any of the traits that can be transferred by a purified DNA preparation.

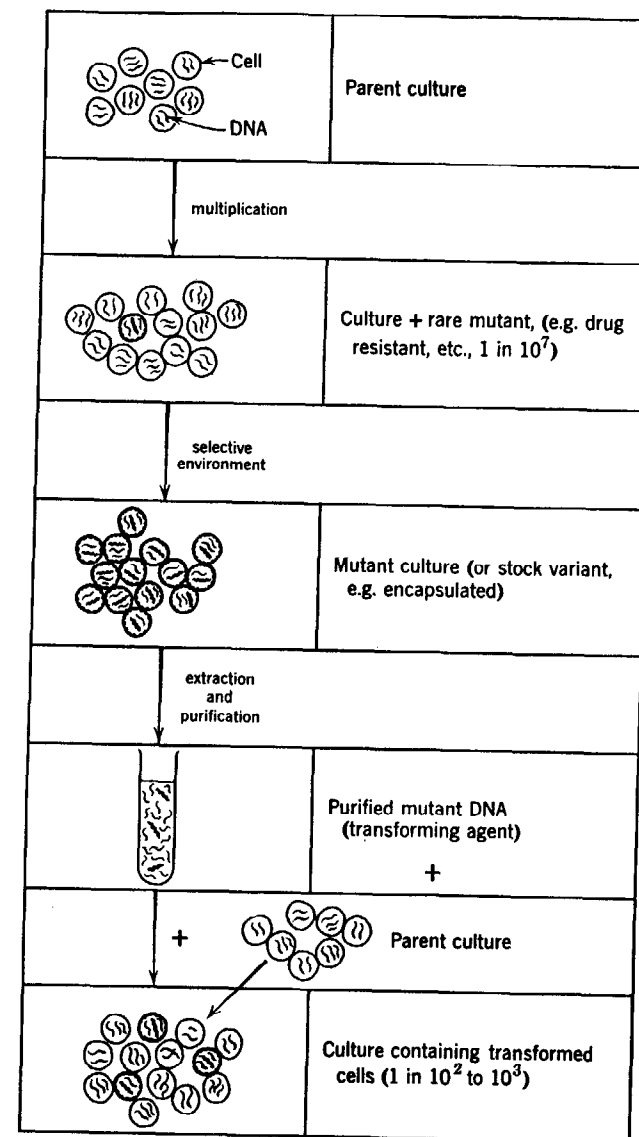


Figure 21. Experimental steps in the transformation of a bacterial culture. Redrawn from R. Hotchkiss, *The Nucleic Acids*, volume 2 (E. Chargaff and J. N. Davidson, editors), Academic Press, 1954.

There are, however, in addition to genetic stability, several other features of transformation that support the notion that the transfer of unit characters by DNA is closely analogous to the normal genetic process occurring during cell division. It has been shown, for example, that the DNA prepared from strains of bacterial cells which have been transformed for two transformable markers can occasionally produce simultaneous transformation for both traits, as though these were linked on the same "chromosomal" fragment. Thus, Marmur and Hotchkiss⁷ have selected, from a strain of streptomycin-resistant pneumococci, mutants that have also undergone a mutation which permits them to ferment mannitol. When DNA preparations from such doubly labeled cells were added to a culture of "wild-type" pneumococci, three distinct types of transformed bacteria could be isolated. The majority were either transformed for streptomycin resistance *or* for mannitol utilization, but some individual cells had clearly been transformed for both characters. The frequency of double transformation was considerably greater than the product of the frequencies of the two single transformations, an observation which is genetically interpretable only in terms of the linked transfer of two unit characters by a single event. This observation, so reminiscent of linkage in the chromosomes of higher organisms, certainly suggests that a single fragment of DNA can contain the information for the elaboration of two distinct physiological systems and that this fragment represents a piece of the normal genetic material of the bacterial cell from which it was derived.

These studies have since been extended by Hotchkiss and his colleagues to cells that contain three transformable markers. They have shown that, during transformation with DNA, these markers may remain linked, be separated, or be reassembled by recombination in a manner completely analogous to that observed with doubly marked cells. In spite of our ignorance of the details of the process, we may conclude with reasonable certainty that the mechanical transfer of genetic information from cell to cell by purified DNA represents a good model for some of the events that occur during conventional gene transfer in dividing cells.

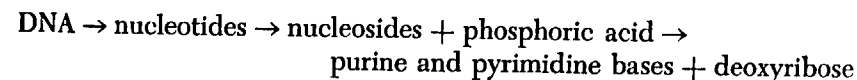
Chemical Structure of DNA

One of the most fascinating recent developments in biochemistry has been the study of the structure of deoxyribonucleic acid, both from the standpoint of its molecular composition and in terms of the

arrangements of its component parts in three dimensions. If DNA is to be established as the substance responsible for the transmission of heredity at the molecular level, we should be able to demonstrate that the morphological behavior of chromosomes can be explained as a function of the structure and metabolism of DNA. Amazingly enough, such a picture has begun to emerge as the result of a series of very skillful experiments and deductions. There are the usual inconsistencies and hopeful extrapolations, but the present outlook is certainly an optimistic one.

In 1948 the conception of DNA as a regular array of repeating tetranucleotide units was seriously shaken by the fundamental studies⁸ of Vischer and Chargaff and of Hotchkiss, who showed by chromatographic techniques that the four heterocyclic bases in DNA are not present in equal amounts and that there are actually more than four such bases in some samples of DNA. Since these studies, the list of naturally occurring purines and pyrimidines in the DNA molecule has grown to seven, and a complete reappraisal of the structural details of DNA has taken place.

By an examination of the products produced from DNA by acid or enzymatic hydrolysis we may deduce that the successive stages of degradation are as follows:*



The chemical structure of these substances is indicated in Figure 22. In most samples of DNA, four heterocyclic bases predominate—the purines adenine and guanine and the pyrimidines thymine and cytosine. However, some samples of DNA (e.g. from wheat germ and the grasses in general) contain 5-methylcytosine in large amounts and, indeed, this derivative of cytosine is found in small quantity in preparations from mammalian tissue as well. The DNA prepared from the "even" members of the T bacteriophages (T2, T4, and T6) contains 5-hydroxymethylcytosine instead of cytosine. An unusual purine, 6-methylaminopurine has been found in the DNA of certain bacteria. For general purposes of discussion, however, we shall take as prototypes only the four common bases.

* The hydrolysis of *ribonucleic acid* (RNA), approximately nine-tenths of which is found in the cytoplasm of cells and one-tenth in the nucleus, yields the same set of products except for *ribose* in place of *deoxyribose*, and *uracil* in place of *thymine* (see Figure 22). The bonds that join the nucleotide residues in RNA are basically similar to those in DNA, but much less is known about its molecular configuration. The presumed role of RNA in the biosynthesis of proteins is discussed in Chapter 10.

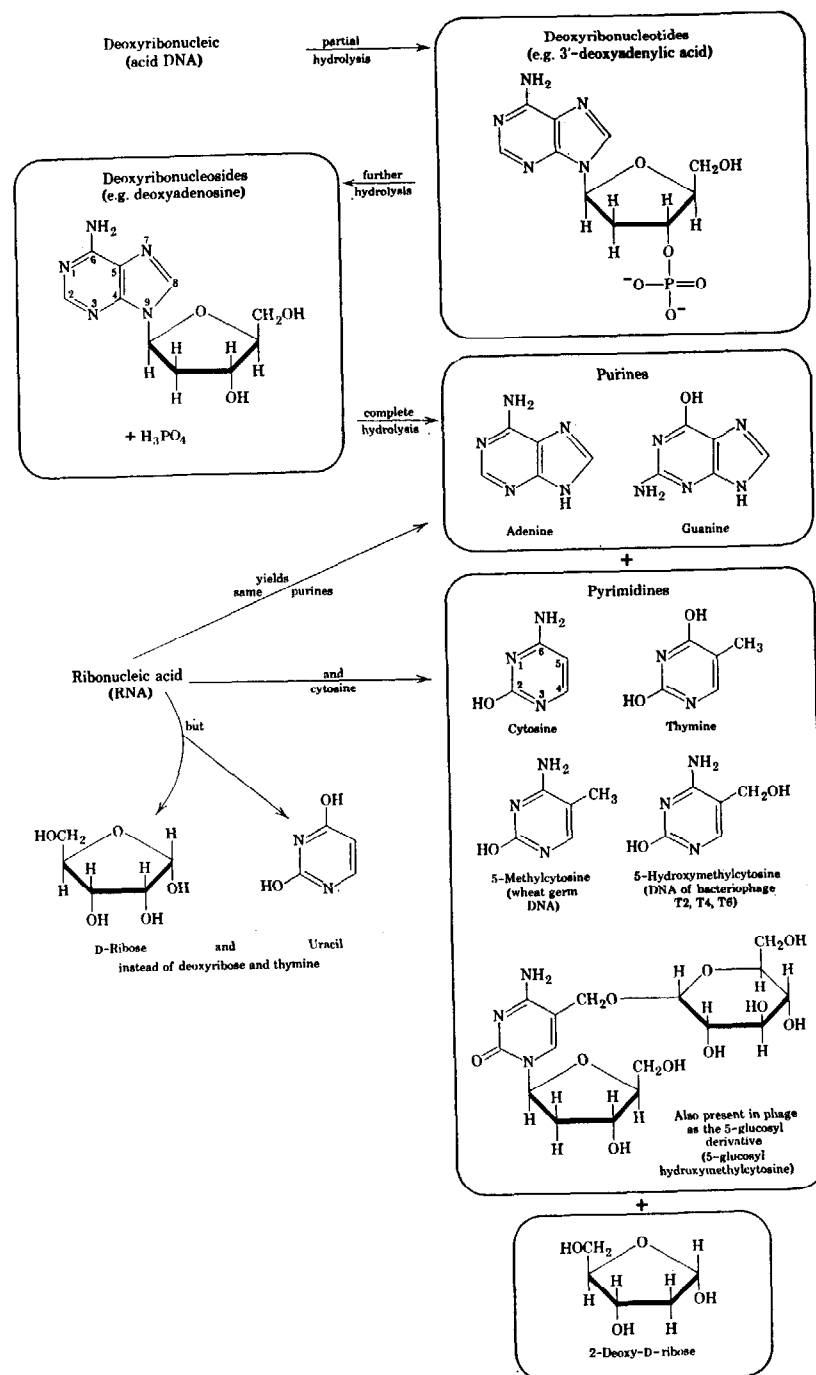
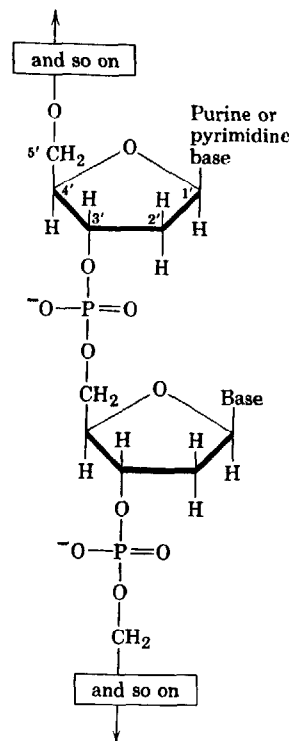


Figure 22. The products of hydrolysis of deoxyribonucleic acid and ribonucleic acid.

The sugar, deoxyribose, together with the phosphate residues, forms the chemical backbone of the DNA molecule. Since deoxyribose has only three hydroxyl groups, and since one of these, at carbon 1, is bound in an N-ribosidic linkage to one of the four heterocyclic bases, the alternating phosphate residues must be doubly esterified with the hydroxyl groups on positions 3 and 5. On the basis of this information, we may already depict the general structure of DNA as shown in Figure 23. To round out the purely chemical part of the DNA structure we need only assign to the various 1 positions of the deoxyribose units the proper choice of purine or pyrimidine base.

This is a very large order. The composition of DNA varies a great deal from source to source and, even more serious, there is now good evidence that most samples of DNA are probably heterogeneous and may represent a mixture of molecules with differing base contents and sequences. In spite of these difficulties, the analytical study of many DNA samples from a host of biological sources has established the following relationships, which hold more or less quantitatively throughout Nature.

Figure 23. The nature of the linkage between individual deoxyribonucleotides in deoxyribonucleic acid. Individual nucleotide residues in the DNA chain are linked through 3'-5' phosphate diester bridges. The carbon atoms of the deoxyribose moiety are numbered 1', 2', etc. The carbon and nitrogen atoms of the purine and pyrimidine rings are numbered 1, 2, 3, etc., without the primes (see Figure 22).



1. There is a stoichiometric equality between the sum of the purines and the sum of the pyrimidines.

2. There is equivalence in the contents of adenine and thymine, and of guanine and cytosine (or the sum of cytosine and its less common derivatives when these are present). Two types of DNA may be recognized, the first in which the sum of adenine plus thymine is larger than the sum of guanine and cytosine (the AT type), and the reverse (the GC type) which occurs mainly in the microorganisms.

3. The content of 6-amino groups equals the content of 6-keto groups.

We have, then, a set of requirements which data on the sequence of bases, now meager but rapidly accumulating, must fit. These

TABLE 1

The distributions of Deoxycytidylic Acid (C) and 5-Methyldeoxycytidylic Acid (M) among the Dinucleotides from Calf Thymus DNA Produced by Digestion with Pancreatic Deoxyribonuclease⁹

C-Dinu- cleotide	Mole Fraction of Digest	M-Dinu- cleotide	Mole Fraction of Digest
C-p-C-p-	1.11	M-p-M-p-	0
C-p-T-p-	0.78	M-p-T-p-	0
T-p-C-p-	2.34	T-p-M-p-	0
C-p-A-p-	0	M-p-A-p-	0
A-p-C-p-	3.22	A-p-M-p-	0
C-p-G-p-	0.75	M-p-G-p-	0
G-p-C-p-	0.12	G-p-M-p-	1.03

The abbreviated formulas for dinucleotides (T-p-C-p- for example) represent the phosphate diesters formed by two mononucleotides, joined in the manner illustrated in Figure 23.

restrictions, in themselves, tell us that DNA molecules must be constructed according to a meaningful plan and that the arrangement of heterocyclic bases must be of great importance in the functional properties of DNA. Other available evidence strongly suggests that the sequence of bases is anything but random. For example, Sinsheimer⁹ has isolated from partial hydrolysates of DNA a large series of dinucleotide fragments, some of which are shown on the left side of Table 1. Some of the sequences listed are much more com-

mon than others. Recently, Shapiro and Chargaff¹⁰ have reported on their similar fractionation studies and concluded that at least 70 per cent of the pyrimidines in various DNA samples occur as oligonucleotide "tracts" containing three or more pyrimidines in a row and that this sort of lumping, because of the three equality relationships listed, must, therefore, also be true of the purines.

This is about as far as we can go at the moment with the critical problem of the sequence of bases. There has been, however, a dramatic advance in connection with the three-dimensional aspects of DNA structure. A large part of this advance stems from the elegant X-ray diffraction studies of M. H. F. Wilkins and his colleagues on DNA from several sources. The X-ray patterns they obtained were all remarkably similar and suggested that there existed some uniform molecular pattern for all deoxyribonucleic acids. Their data were consistent with the presence in DNA of two or more polynucleotide chains arranged in a helical structure and permitted the calculation of the rough dimensions of the helix.

Employing the data of Wilkins and his collaborators,¹¹ and taking into consideration the various restrictions on base pairing and stoichiometry, J. D. Watson and F. H. Crick¹² were able to construct a model which accommodates most of the experimental facts. This model, in spite of slight stereochemical shortcomings which have since been adjusted by Wilkins and his colleagues (Figure 24), has been a tremendously valuable stimulus to the field of nucleic acid chemistry as well as to genetics. The Watson-Crick model was based, in addition to the analytical data on base content and the X-ray results, on the fact that titration studies had indicated that the polynucleotide chains in the DNA molecule were joined together through hydrogen bonding between the base residues. The critical assumptions made by Watson and Crick were that the number of chains in the molecule was two, and that there was a specific sort of base pairing which made each pair symmetrical and equivalent in relation to the cross-linking of the two sugar-phosphate backbones. The base pairs required by the model, adenine-thymine and guanine-cytosine, are shown in Figure 25, joined together through hydrogen bonds. The formulas shown in the figure are drawn from the work of Linus Pauling and Robert Corey,¹³ who have considered the DNA double-helix model in detail in relation to the crystal structure data for purines and pyrimidines. Their results, which show that three hydrogen bonds can be formed between cytosine and guanine, indicate that an even greater specificity of base pairing is inherent in the DNA model than was suggested by the considerations of Watson

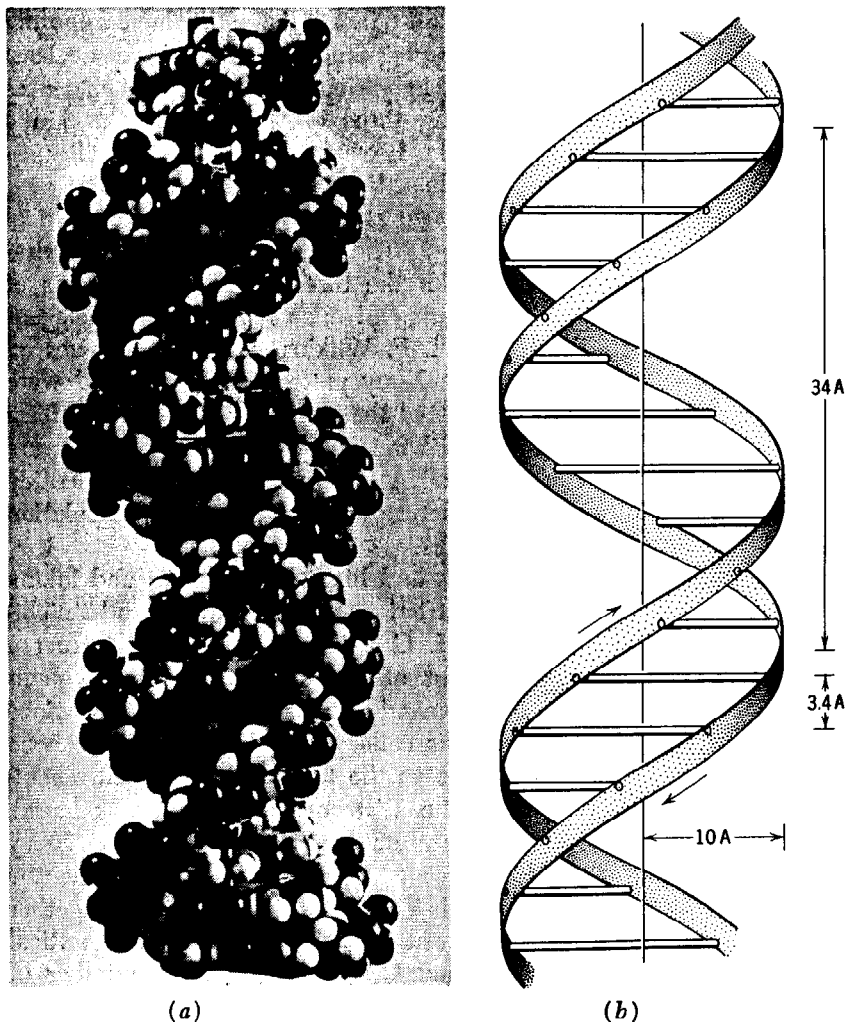


Figure 24. Photograph of a molecular model of deoxyribonucleic acid (through the courtesy of Dr. M. H. F. Wilkins). A schematic drawing of this two-stranded structure is shown on the right, together with certain of its dimensions.

and Crick, who assumed that two hydrogen bonds are formed between each base pair.

The model has the interesting feature that the two strands of polynucleotide are complementary to one another and that the arrangement of bases on one strand fixes the arrangement on the

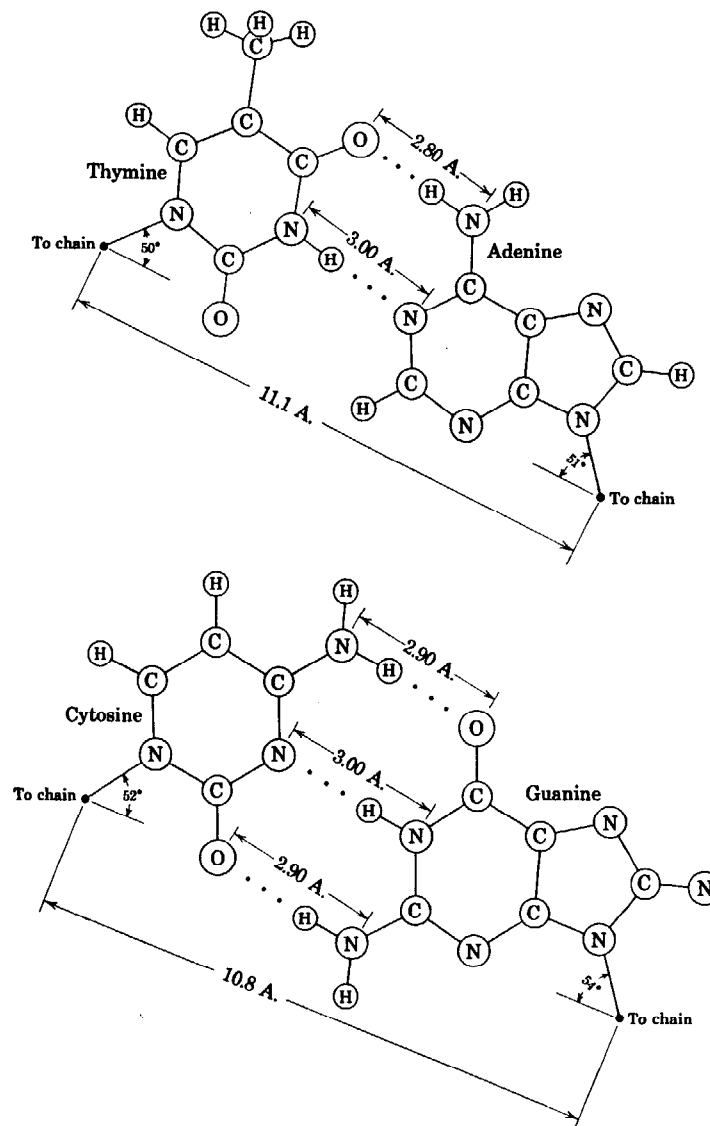


Figure 25. The pairing of adenine and thymine (top) and cytosine and guanine (bottom) by means of hydrogen bonding. Redrawn after L. Pauling and R. B. Corey, *Arch. Biochem. Biophys.*, **65**, 164 (1956).

other. Thus, if the bases along one strand are arranged in the order A-G-G-T-C-A, the opposite bases on the complementary strand will be T-C-C-A-G-T. This fact has interesting implications in connection with the process of genetic replication, which we shall shortly discuss.

The Watson-Crick model has received considerable support and is certainly a splendid working hypothesis. We must remember, however, that it is only a model and therefore examine both the pros and the cons.

In support of the double helix can be offered the studies of Thomas¹⁴ and of Schumaker, Richards, and Schachman.¹⁵ Both have concluded, from investigations of the changes in light scattering and viscosity during partial enzymatic digestion, that the bulk of the DNA molecule must be in the form of double chains which are linked together. They have calculated that, if the molecule were a single chain, the mean molecular weight of the fragments of DNA produced should have been considerably smaller than observed and that the facts correspond well with a system of rather stable cross-linkages between two polynucleotide strands.

Some of the most impressive support for a double-stranded, complementary structure comes from the very recent work of Meselson and Stahl¹⁶ using an ultracentrifugal method developed by Vinograd¹⁷ and his group at the California Institute of Technology. Vinograd has capitalized on the fact that strong salt solutions (e.g. of cesium chloride) may be forced into a density gradient when exposed to high centrifugal fields.

A macromolecular substance like DNA dissolved in the salt solution will seek its own density in the gradient and will form a relatively sharp line in the centrifugal cell; the line may be detected by its strong absorption of ultraviolet light. If two species of DNA that differ in density are present, these will separate and form two distinguishable zones (Figure 26). In this way it is possible to estimate, for example, the relative proportions of normal DNA and of DNA containing the rather dense pyrimidine, bromouracil, in a mixture of the two.

Meselson and Stahl grew *Escherichia coli* in a nutrient medium containing nitrogen of mass 15 exclusively and thus obtained a population of organisms which were fully labeled with this isotope. They then transferred cells, taken during the logarithmic phase of growth, to a medium containing N¹⁴ precursors. Samples were taken when the population had doubled and again after a second doubling. Deoxyribonucleic acid was prepared from the original fully labeled

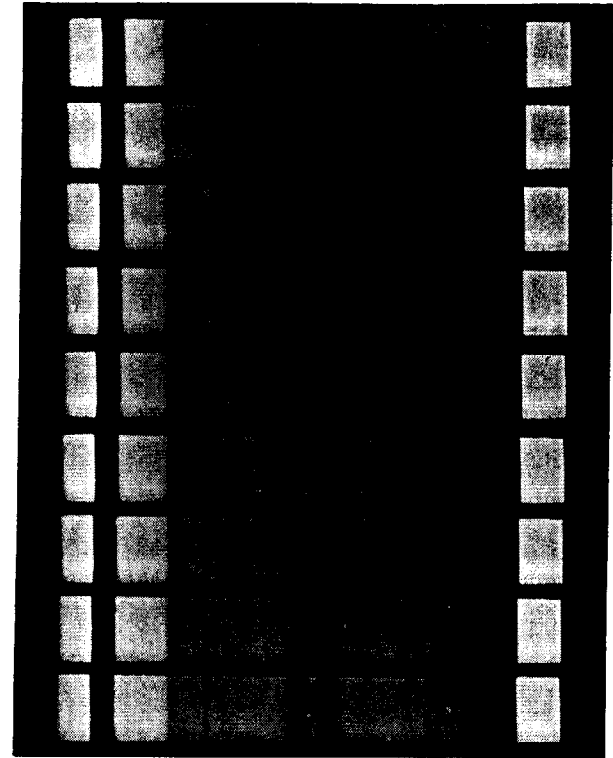


Figure 26. The density gradient centrifugation of a mixture of N¹⁵- and N¹⁴-labeled deoxyribonucleic acids from *E. coli*. Exposures were taken every 128 minutes. Equilibrium has been attained by the time of the eighth or ninth photograph. The molecular weight of the DNA is approximately 10,000,000. From M. Meselson and F. W. Stahl, *Proc. Natl. Acad. Sci. U.S.*, **44**, 671 (1958).

cells and from the two batches of cells that had undergone multiplication. Each of these preparations was then dissolved in strong cesium chloride solution, having a density in the neighborhood of the apparent density of DNA, and centrifuged until density equilibrium had been attained in the ultracentrifuge cell. Reproductions of the distribution patterns obtained are shown in Figure 27. The DNA prepared from the fully labeled bacteria gives a single band corresponding to the density to be expected for N¹⁵ material. After one generation this band has disappeared and has been replaced with a band having a density to be expected for equal amounts of N¹⁴ and N¹⁵. Finally, after two generations, *two* bands are observed, one corresponding to 100 per cent N¹⁴-DNA and one to the 50:50

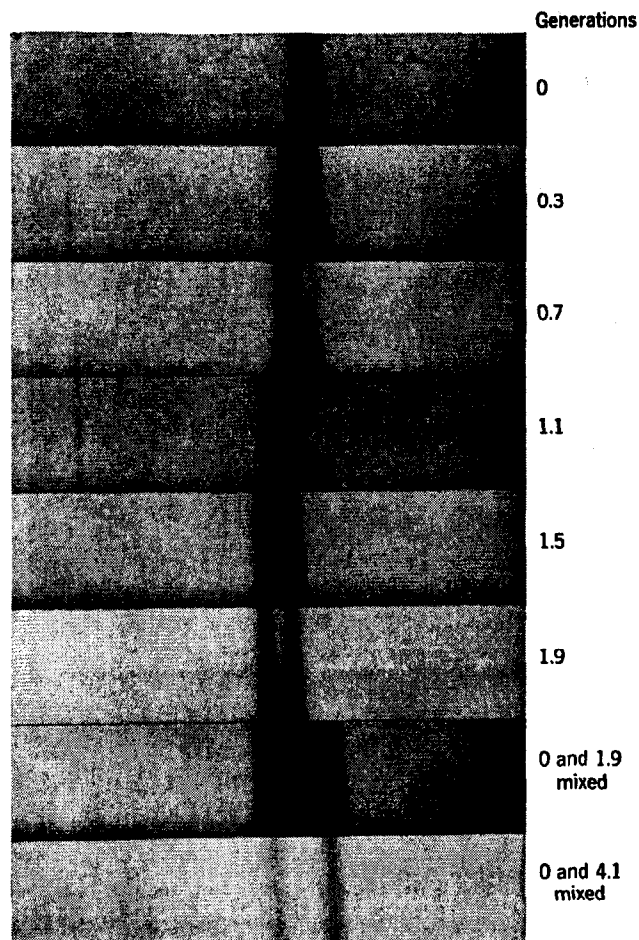


Figure 27. Ultraviolet absorption photographs showing the concentration of DNA by density gradient centrifugation of lysates of bacteria sampled at various times after the addition of N^{14} substrates to a growing N^{15} -labeled culture. The length of exposure to N^{14} substrates is indicated in terms of generations. After one generation (fourth frame from the top) the density of the bacterial DNA lies halfway between that of the original N^{15} -labeled material and that of material containing essentially only N^{14} (the left component in the bottom frame). Taken from the work of M. Meselson and F. W. Stahl, *Proc. Soc. Natl. Acad. Sci. U.S.*, **44**, 671 (1958).

N^{14} - N^{15} mixture. In experiments in which cells were permitted to multiply through further generations, the mixed band decreased in amount and was replaced by a greater and greater proportion of the band corresponding to N^{14} -DNA. If replication of DNA had occurred by a process which involved intermediate degradation into smaller pieces with subsequent reutilization of these pieces for the synthesis of new DNA molecules, density of the first-generation molecules could not have corresponded so sharply to a specific equimolar mixture of "light" and "heavy" chains. We would expect, rather, a broad distribution of densities, under such circumstances, quite unlike the results that were actually obtained.

These results are in good agreement with the predictions of the Watson-Crick model. As shown schematically in Figure 28, each strand of the DNA double helix might be visualized as a template on which a new strand is fashioned during replication. If the precursor nucleotides are unlabeled, the two "daughter" helices produced during a single replication would be expected to be half-labeled. After a second replication two strands would be labeled and two unlabeled. Base pairing would assure accurate assembly of complementary strands, and genetic constancy would thus be maintained.

The scheme of base pairing suggested by Watson and Crick has recently received strong support from the enzymatic studies of Arthur Kornberg and his colleagues¹⁸ on the *in vitro* synthesis of DNA-like molecules. They reported, in 1957, that an enzyme had been purified from extracts of *E. coli* which could catalyze the condensation of deoxyribonucleoside triphosphates to form large polydeoxynucleotide chains with the liberation of pyrophosphate in the presence of a DNA "primer." In subsequent studies the enzyme has been purified 4000-fold over the crude extract, and a number of important properties of the system have been delineated. The most dramatic finding has been that the composition of the newly synthesized polymer (having a molecular weight of the order of 5,000,000) exhibits the same purine and pyrimidine base composition as that of the primer, in spite of a constant concentration of each triphosphate added. For example, the addition of DNA from *Aerobacter aerogenes*, having a ratio of adenine plus thymine to guanine plus cytosine of 0.82, induced the formation of "DNA" with a ratio of 0.99. With DNA from T2 coliphage, having a ratio of 1.91, the new material showed a value of 1.98. In keeping with the base pairing of the Watson-Crick model, deoxyuridine triphosphate was incorporated into DNA only in place of thymidine triphosphate and deoxyinosine triphosphate only in place of deoxyguanosine triphosphate. This major break-

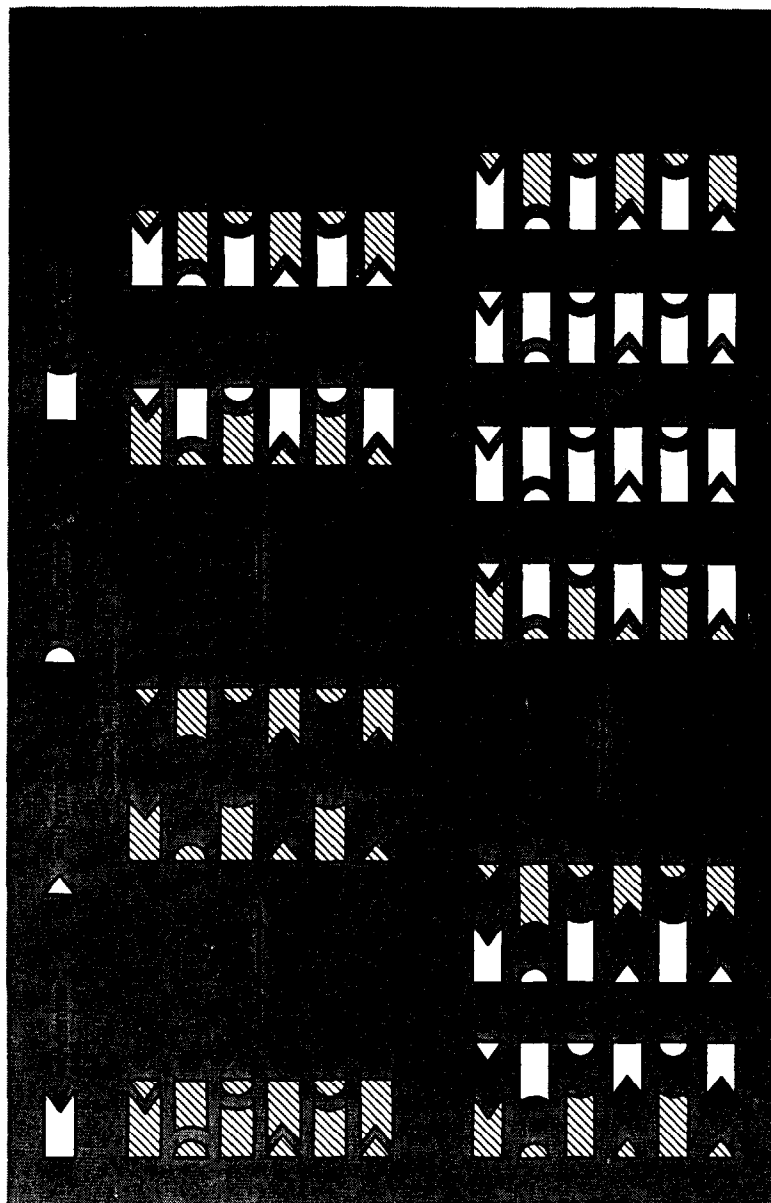


Figure 28. A schematic representation of how the replication of deoxyribonucleic acid might occur. The base sequence in a newly formed strand is determined by the complementarity between adenine and thymine and between guanine and cytosine. A process such as that shown here would be compatible with the results obtained by Meselson and Stahl, some of which are shown in Figure 27.

through, in addition to supplying evidence bearing on DNA structure, may also be a key to the ultimate synthesis of polydeoxynucleotides having specific biological activity.

The studies we have discussed deal with genetic material at the molecular level. To make the results meaningful to biology, they must be shown to apply to, or at least to be consistent with, the behavior of DNA at the level of the organized chromosome. The extrapolation is enormous and consequently entails a good deal of faith. We can determine, for example, that the DNA in a single haploid cell (of *Lilium*) is approximately 53×10^{-12} grams. If we assume that this DNA is present as a single, long double helix, it can be calculated that the length of this coil would be 1.5×10^7 microns, or 15 meters, and that the coiled structure would make 4.4×10^9 turns around the screw axis. The mechanical problems of unwinding such a coil during the replication process are of a magnitude to make a picture as simple as this one of no serious value.

In spite of the enormous gaps in knowledge between DNA structure and chromosome organization, recent experiments by Taylor, Woods, and Hughes¹⁹ on chromosome replication make it clear that some hypothesis must be arrived at which will accommodate both chemical and cytological information. Taylor et al. have performed the equivalent of the Meselson-Stahl experiment at the cellular level. Using a technique similar to that employed by Howard and Pelc (page 41), they have studied the distribution of label between daughter cells in successive generations. Their radioautographic technique was made considerably sharper by the use of tritium-labeled thymidine rather than radioactive phosphate as a labeled precursor. Since tritium emits beta particles of rather low energy in comparison with phosphorus, film blackening is more highly localized and the resulting autoradiographs have a high degree of correspondence with cellular structure as observed by microscopic examination.

Vicia faba seedlings were grown in a medium containing tritiated thymidine. During this incubation most of the twelve chromosomes became labeled as evidenced by the correspondence between the location of silver grains in the photographic emulsion and the microscopically visible chromosomes. Roots at this stage were then transferred to nonradioactive thymidine solutions containing colchicine and incubated for varying lengths of time. In the presence of colchicine, *Vicia* chromosomes contract to the metaphase condition and sister chromatids (the two separate strands of the chromosome duplex) are spread apart and are easily observable. Since colchicine

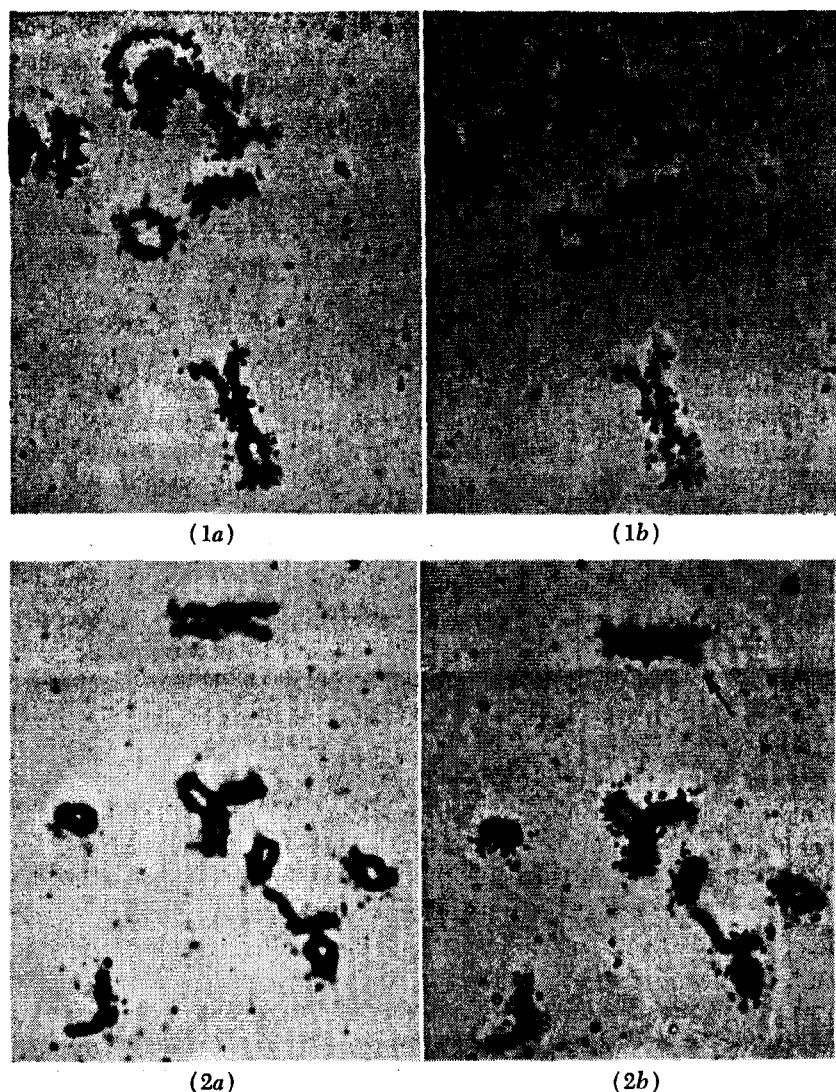


Figure 29. *Vicia faba* chromosomes. Upper half, chromosomes in metaphase, at the first division after labeling with radioactive thymidine: (1a) chromatids spread apart but attached at the centromere; (1b) silver grains in the photographic emulsion above the chromosomes. Lower half, chromosomes after labeling, showing replication in the absence of labeled precursor: (2a) chromosomes from a cell containing 24 chromosomes, with chromatids spread but attached at the centromere; (2b) silver grains in the emulsion above the chromosomes in 2a. As described in the text, these results suggest that the formation of new chromatids takes place at the expense of low molecular weight precursors and does not involve the degradation and reutilization of pre-existing deoxyribonucleic acid.

has the effect of preventing anaphase movements and the subsequent separation of daughter cells, but does not stop chromosomes from duplicating, a count of the number of chromosomal pairs in each cell gives a measure of the number of rounds of replication. Thus, cells with 24 and 48 chromosomes have replicated once and twice, respectively.

Taylor and his colleagues observed that in cells with 24 chromosomes only one of the sister chromatids in each metaphase pair was labeled. It was evident that the pool of labeled precursor had been rapidly depleted in the cells after removal from the initial solution and that replication of chromosomes to yield 24 had involved the use of the nonradioactive thymidine.

In cells with 48 chromosomes, analysis of all 48 was not possible. However, of those that were sufficiently separated for clear observation, approximately half contained one labeled and one unlabeled sister chromatid and half were completely unlabeled. In an occasional "second generation" pair of chromosomes, individual sister chromatids were labeled for only part of their length (see arrow in Figure 29), and in such cases the corresponding portion of the other chromatid showed radioactivity. This behavior is to be expected for crossing over and resembles the cytological observations that have been made for the chromosomes of other species (e.g. the giant salivary gland chromosomes of *Drosophila*).

The results are interpretable in terms of a scheme such as shown in Figure 30. This diagram proposes the presence in each chromosome of two morphological halves held together at the centromere

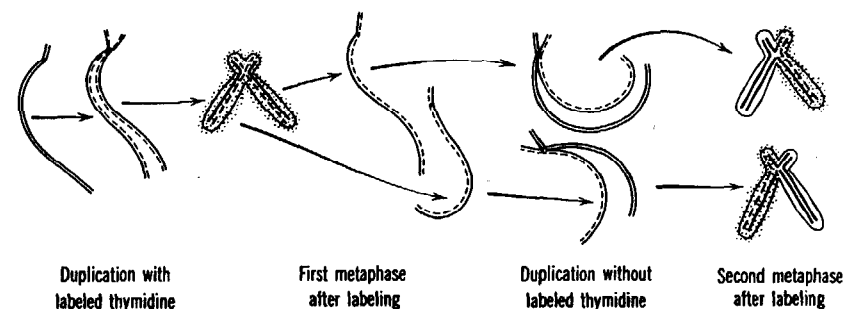


Figure 30. A schematic representation of the photographs in Figure 29. Solid lines indicate unlabeled units in the chromosome and dashed lines represent radioactive units formed during replication. The dots represent grains in the photographic emulsion caused by radioactive decay of the tritiated thymidine.

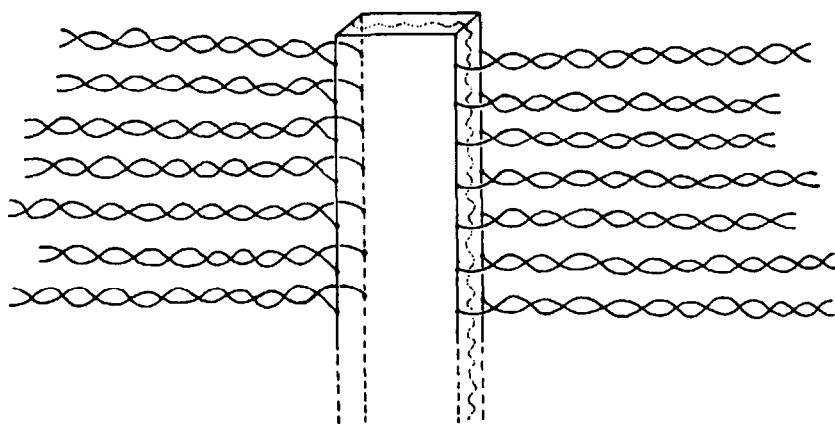


Figure 31. A schematic diagram showing how the chromosome might be organized to account for the results summarized in Figure 29. The chromosome is depicted as consisting of a central core divided into two halves. Each half is attached to one strand of a large number of DNA double helices. Redrawn from J. H. Taylor, *Am. Naturalist*, 91, 209 (1957).

and containing two equivalent strands. During the initial labeling period each chromatid becomes radioactive, and metaphase figures in cells with twelve chromosomes are uniformly associated with blackening of the emulsion. After a second replication each chromosome pair is still radioactive, but all the radioactivity is associated with the original parent strand and the newly formed strand is unlabeled.*

Realizing the unlikelihood of a double helix running the full length of the chromosome, Taylor, Woods, and Hughes have proposed an interesting model which accommodates both their own radioactive data at the cellular level and the theoretical implications of the double helix. This model (see Figure 31) visualizes the chromosome as composed of two halves, each half consisting of a central core to which is attached one strand of each of a multitude of double helices. When a cell enters the division cycle, the halves separate from one

* L. F. La Cour and C. R. Pele have recently repeated these experiments in the *absence* of colchicine and report that radioactivity was distributed between *both* chromatids. No free labeled precursors were present during the period of replication. These results, which are in direct contrast to those of Taylor, Woods, and Hughes, reopen the question of whether replication is, or is not, a completely "conservative" mechanism in which a chromosomal strand, once formed, is retained in the nucleus as an essentially permanent chemical structure.

another, each half carrying with it one of the complementary polynucleotide chains. Taylor and his colleagues point out that the model introduces a new dimension for the interpretation of crossing over. Thus, conventional crossover might occur by exchange along the core, whereas various minor recombinations and gene conversions could occur by interaction or exchange among the side chains of DNA.

Summing up, we may say that a good case can be made for the presence of a set of intertwined helical coils of polynucleotide in DNA. The model proposed by Watson and Crick is, at present, the most satisfactory. Although J. Donohue²⁰ has shown that certain other arrangements of base pairs can be made which would permit the construction of double-helix models, these do not agree with the X-ray patterns that have been observed for DNA, and the pairing of adenine with thymine and of guanine with cytosine seems most probable at the moment. The complementarity of the two strands forms the basis of an intriguing hypothesis for replication of genetic material, supported by experimental evidence at both the molecular and the cellular level.

Let us now consider briefly some of the bits of evidence that do *not* fit entirely with the Watson-Crick hypothesis. These are not as numerous as the supporting evidence, possibly because scientists, like other citizens, are also attracted to bandwagons.

It has been suggested that the replication of DNA based solely on complementarity of polynucleotide chains might be difficult to reconcile with some of the analytical values for nucleotide abundance. Sinsheimer⁹ has isolated a series of dinucleotides from calf thymus DNA (see Table 1) which includes homologous dinucleotides of cytosine and of 5-methylcytosine. The results indicate a considerable difference in the distribution of cytosine and 5-methylcytosine; that is, substitutions do not appear to be random. Although still somewhat preliminary, these findings suggest that simple complementarity might not be an adequate basis for the distribution. There is no obvious stereochemical reason for discrimination against 5-methylcytosine in positions usually occupied by cytosine, and Sinsheimer justifiably recommends that we keep our eyes open for ancillary pathways of nucleotide metabolism or assembly that might explain such deviations from simplicity.

Sinsheimer has also pointed out another more serious objection to simple complementary replication. In T2 bacteriophage, in which cytosine is completely replaced by 5-hydroxymethylcytosine, about 77 per cent of the latter compound occurs conjugated with glucose

(see Figure 22). Now various studies (which we shall discuss in the next chapter) have indicated that the DNA of bacteriophage T2 is composed of one large piece comprising about 40 per cent of the total and a number of smaller pieces. All the unconjugated 5-hydroxymethylcytosine is found in the large piece; that is, only about 60 per cent of the large piece is substituted, whereas all the smaller pieces contain glucose. As stated by Sinsheimer, "Since these components of T2 DNA are all replicated in the same cell at the same time, the limited glucose substitution of the large piece cannot reasonably be ascribed to some metabolic limitation. Hence it would appear (to put the matter mechanistically) that when the replication of the large piece takes place, the machinery knows when to put in glucose-substituted 5-hydroxymethylcytosine or when to insert unsubstituted 5-hydroxymethylcytosine, for the glucose content remains unchanged through many generations." These observations could be explained on the basis that 5-hydroxymethylcytosine and its glucose-substituted derivative present, to the assembly line, different stereochemical features which would preclude error in the determination of sequence. The cautions advanced by Sinsheimer concerning simple complementarity must obviously be kept well in mind, both as brakes on overenthusiasm and for their heuristic value.

Some of the most puzzling evidence contrary to the Watson-Crick replication hypothesis comes from the work of G. Stent and his co-workers on "suicide" in P³²-labeled phage as the result of radioactive decay. These experiments are not easily described, however, without some preliminary consideration of the natural history of bacteriophage; consequently they will be discussed in the following chapter.

Polynucleotide "Codes"

No discussion of DNA in relation to genetics would be complete without some mention of the efforts that have been made to discover a "code" which might permit the translation of base sequence into protein structure. To participate in this game we need have no special knowledge of protein chemistry beyond the following two basic facts: first, proteins consist, for the most part, of polypeptide chains made up of approximately twenty amino acids (*exactly* twenty in codes to date) arranged in a specific sequence for each protein species; and second, the known protein and polypeptide sequences have been assembled in several review articles which are

... B C A, C D D, A B A, B D C, . . .
or
... B, C A C, D D A, B A B, D C . . .

Figure 32. Linear arrangements of four purine or pyrimidine bases, A,B,C and D, in which triplets, each corresponding to a specific amino acid residue, are separated by commas. The problem is how to read the code if the commas are rubbed out.

easily available in most libraries. Some mathematical facility and considerable imagination are prerequisites.

We shall consider only one such code here, the so-called "comma-less code" of Crick, Griffith, and Orgel.²¹ It makes a good example because experimental data have not yet appeared which can rule out this code, and the basic cryptographic approach to "how a sequence of four things [nucleotides] determines a sequence of twenty things [amino acids]" is elegantly illustrated in their paper.

Since there are only four (common) bases, a code founded on correspondence between base *pairs* and amino acid residues would not suffice, only $4 \times 4 = 16$ such pairs being possible. Base *triplets*, on the other hand, are too numerous ($4 \times 4 \times 4 = 64$) without the introduction of restrictions. As one approach to restricting the possibilities, Crick, Griffith, and Orgel have chosen to consider the case of "nonoverlapping codes." An illustration from their discussion states the problem clearly. Letting A, B, C, and D represent the four bases, we may divide a polynucleotide chain of such bases into nonoverlapping triplets, each triplet corresponding to an amino acid residue (Figure 32). "If the ends of the chain are not available, this can be done in more than one way as shown. The problem is how to read the code if the commas are rubbed out, that is, a comma-less code."

The ground rules for this particular coding puzzle were as follows: (1) Some of the 64 triplets make sense, and some "make nonsense"; (2) all the possible sequences of amino acids may occur, and at every point in the string of bases we can only read sense in the correct way. To put ground rule 2 in other words, two triplets that make sense can be put side by side, but overlapping triplets so formed must always be nonsense. Thus, in the sequence ABCDAA, if ABC and DAA make sense, BCD and CDA must always be nonsense.

To prove that 20 is the maximum number of amino acids that can be coded in this manner is quite straightforward. Thus we may

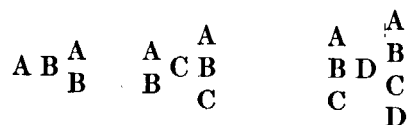


Figure 33. One solution to the problem of how twenty amino acids can be coded by four heterocyclic bases in a "comma-less" code.

eliminate AAA, BBB, etc., since, if in a sequence AAAAAA the triplet AAA corresponds to amino acid *a*, the sequence of six may be misinterpreted by associating acid *a* with the second to fourth, or third to fifth letters. This reduces the 64 possibilities to 60. These 60 may be grouped into twenty sets of three, each set being cyclic permutations of one another. If BCA stands for amino acid *b*, for example, then BCABCA stands for *bb*, and CAB and ABC must be ruled out as nonsense. Since only one triplet can therefore be chosen from each cyclic set, we arrive at the "magic number" of 20.

Having shown that no more than twenty amino acids can be coded by four bases in this system, the authors list a number of solutions, including that shown in Figure 33. The reader may satisfy himself that any two triplets of this set may be placed next to each other without producing overlapping triplets which belong to the set. In all, 288 different solutions were found which conformed to the rules.

If the structure of DNA does, indeed, have some direct relationship to genetic information, we must undoubtedly take into account many structural factors in addition to base sequence, particularly those having to do with the geometrical arrangement of the polynucleotide chain. The business of constructing codes that might have some bearing on the problem of information transfer from DNA to protein is, therefore, mainly of theoretical value at the moment. In time, however, the study of DNA-protein cryptography will undoubtedly become a much more active field. Protein chemists are now beginning to accumulate data relating to the chemical consequences of mutation, and there should soon be a large amount of information for consideration. These data will be, for the most part, very indirect. For example, the single-gene change which leads to the formation of hemoglobin C instead of the normal hemoglobin A is reflected in the replacement of a particular glutamic acid residue by a residue of lysine (Chapter 8). The same replacement has occurred, during evolution, in the ribonuclease molecule where a lysine residue in bovine ribonuclease is replaced by glutamic acid

in the sheep protein (Chapter 7). In the first case we can definitely invoke mutation as a causative factor. In the second, mutation is implied but, in view of the impossibility of crossing cows with sheep, a cause and effect relationship cannot be proved. As further examples of this sort accumulate, we shall, perhaps, be able to decide whether there exists some uniform pattern of correspondence between genetic composition and protein structure.

REFERENCES

1. A. Boivin, R. Vendrely, and C. Vendrely, *Compt. rend.*, **226**, 1061 (1948).
2. A. E. Mirsky, in *Genetics in the 20th Century* (L. C. Dunn, editor), Macmillan, New York (1951).
3. S. Zamenhof, R. DeGiovanni, and K. J. Rich, *Bacteriology*, **71**, 60 (1956).
4. A. Howard and S. R. Pelc, *Exptl. Cell Research*, **2**, 178 (1951).
5. O. T. Avery, C. M. MacLeod, and M. McCarty, *J. Exptl. Med.*, **79**, 137 (1944).
6. These studies are reviewed by R. D. Hotchkiss in the book, *The Nucleic Acids*, volume 2 (E. Chargaff and J. N. Davidson, editors), Academic Press, New York, 1955.
7. J. Marmur and R. D. Hotchkiss, *J. Biol. Chem.*, **214**, 383 (1955).
8. Discussed by E. Chargaff in *The Nucleic Acids*, volume 1 (E. Chargaff and J. N. Davidson, editors), Academic Press, New York, 1955.
9. R. L. Sinsheimer, *Science*, **125**, 1123 (1957).
10. H. S. Shapiro and E. Chargaff, *Biochim. et Biophys. Acta*, **26**, 608 (1957).
11. M. H. F. Wilkins, A. R. Stokes, and H. R. Wilson, *Nature*, **171**, 738 (1953).
12. J. D. Watson and F. H. C. Crick, in *Viruses, Cold Spring Harbor Symposia Quant. Biology*, **18**, (1953).
13. L. Pauling and R. B. Corey, *Arch. Biochem. Biophys.*, **65**, 164 (1956).
14. C. A. Thomas, *J. Am. Chem. Soc.*, **78**, 1861 (1956).
15. V. N. Schumaker, E. G. Richards, and H. K. Schachman, *J. Am. Chem. Soc.*, **78**, 4230 (1956).
16. M. Meselson and F. W. Stahl, *Proc. Natl. Acad. Sci. U.S.*, **44**, 671 (1958).
17. M. S. Meselson, F. W. Stahl, and J. Vinograd, *Proc. Natl. Acad. Sci. U.S.*, **43**, 581 (1957).
18. J. Adler, M. J. Bessman, I. R. Lehman, H. K. Schachman, E. S. Simms, and A. Kornberg, *Federation Proc.*, **16**, 153 (1957). See also *Federation Proc.*, **17**, 178 (1958).
19. J. H. Taylor, P. S. Woods, and W. L. Hughes, *Proc. Natl. Acad. Sci. U.S.*, **43**, 122 (1957).
20. J. Donohue, in *Molecular Structure and Biological Specificity*, (L. Pauling and H. Itano, editors), American Institute of Biological Sciences, Publication 2, Washington, D. C., 1957.
21. F. H. Crick, J. S. Griffith, and L. E. Orgel, *Proc. Natl. Acad. Sci. U.S.*, **43**, 416 (1957).

SUGGESTIONS FOR FURTHER READING

- Biochemical Society, *The Structure of Nucleic Acids and Their Role in Protein Synthesis*, Cambridge University Press, Cambridge, England, 1957.
- The Chemical Basis of Heredity* (W. D. McElroy and B. Glass, editors), Johns Hopkins Press, Baltimore, 1957.
- Crick, F. H. C., in *Sci. American*, **197**, No. 3, 188 (1957).
- Overend, W. G., and A. R. Peacocke, *Endeavour*, **16**, No. 62, 90 (1957).
- Taylor, J. H., "The Time and Mode of Duplication of Chromosomes," *Am. Naturalist*, **91**, 209 (1957).